

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



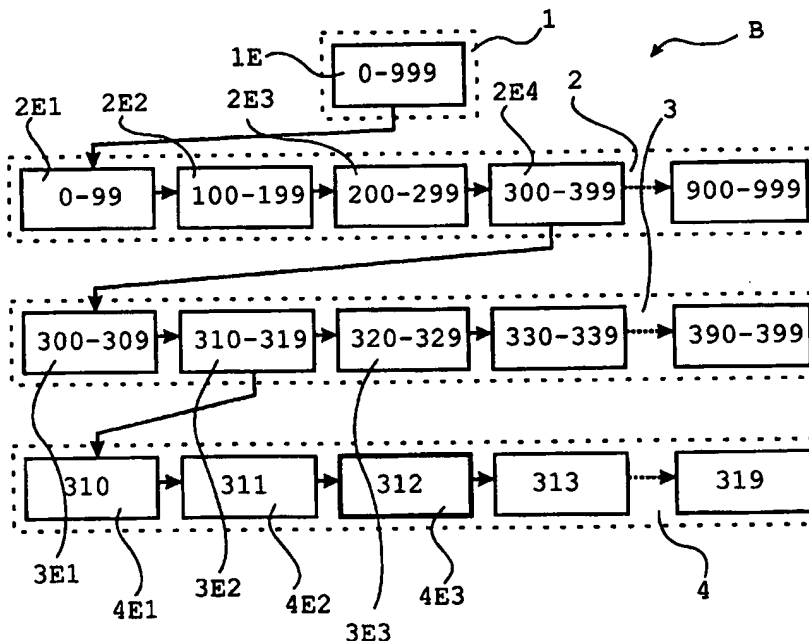
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 17/30		A2	(11) International Publication Number: WO 99/05617
			(43) International Publication Date: 4 February 1999 (04.02.99)
(21) International Application Number: PCT/SE98/01333			(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 7 July 1998 (07.07.98)			
(30) Priority Data: 9702761-9 21 July 1997 (21.07.97) SE			
(71) Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (publ) [SE/SE]; S-126 25 Stockholm (SE).			
(72) Inventor: RONSTRÖM, Ulf, Mikael; Örnbacken 14, S-126 51 Hägersten (SE).			
(74) Agent: LINDBLOM, Erik, J.; Flotthamn, S-150 23 Enhörna (SE).			Published <i>Without international search report and to be republished upon receipt of that report.</i>

(54) Title: STRUCTURE FOR A DATA-BASE

(57) Abstract

The present invention relates to a structure for the storage of data-information within a data-base, which information is structured into a hierarchy of nodes and which hierarchy consists in a tree-structure. This tree-structure comprises a root-node (1), possibly one or several intermediate nodes (2, 3) and one or more leaf-nodes (4), wherein the root-node (1) comprises one or more root-elements (1E), wherein respective intermediate nodes (2, 3) include one or more intermediate elements (2E1, 2E2, 2E3, ...), and wherein respective leaf-nodes (4) include one or more leaf-elements (4W1, 4E2, 4E3, ...). Respective elements include the address information necessary to hold the tree-structure together and enable searches to be performed in the tree and data-information. The order of the elements within a



node is a logical order concerning the data-information, such as a numerical or alphabetical order, where the data-information within respective elements constitutes the difference in data-information in relation to the data-information within the preceding element in the hierarchy. The data-information within the root-element constitutes a start-value. This start-value is the actual data-information, and the differential information within the remaining elements is made up of a data-word that represents the actual difference of the data-information in relation to the data-information belonging to the preceding element, regardless of where the difference is within the data-information.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5

10 **TITLE OF INVENTION:** Structure for a data-base.

FIELD OF INVENTION

15 The present invention relates to a structure for the storage of data-information within a data-base.

The structure is suitable where data-information is structured in a node hierarchy that comprises a B-tree. This tree includes root-node, possibly one or more levels of intermediate nodes, and one or more leaf-nodes.

20

The root-node includes a root-element, respective intermediate nodes include one or more intermediate elements, and respective leaf-nodes include one or more leaf-elements.

25

The B-tree is built-up by linking a first intermediate element within an intermediate node to the root-element or to an intermediate element within an intermediate node placed at a higher level in the hierarchy. Remaining intermediate elements within the intermediate node are linked to the first intermediate element.

30

Similarly, a first leaf-element within respective leaf-nodes is linked either to the root-element or to an intermediate element of one of the intermediate nodes. Remaining leaf-elements in a leaf-node are linked to the first leaf-element.

35

Respective elements include address-information for the links and data-information.

5 The order of the elements within a node is a logical order with respect to data-information, such as an alphabetical or numerical order.

10 The data-information in respective elements constitutes the difference in data-information in relation to the data-information belonging to preceding elements in the hierarchy.

BACKGROUND OF THE INVENTION

15 It has long been known to use different data structures to store data within a data-base. These structures afford various advantages with respect to used storage space and the possibility of fast-searching, for instance.

20 A known structure is based on a B-tree, thereby enabling the logical distribution of elements in a table within a data-base.

25 An element may include an internal address in a data-base that is coupled to an external key, such as a telephone number, for instance. When inserting a new telephone number in a table, a calculation is made as to the element in which the internal address of the telephone number shall be placed, by means of a logic order.

30 The function of the B-tree is to distribute the addresses in the table, so that said addresses are distributed in a manner which will enable them to be easily found.

35 The B-tree is used essentially to create an ordered and simple structure.

Each tree commences with a root-node which points to further nodes in which a search for a desired node can be carried out

on the basis of certain criteria. These nodes will often include a key or an index.

Http addresses, file names, time stamps, telephone numbers,
5 etc., are examples of keys stored in a B-tree.

It should also be mentioned that the aforesaid nodes are nodes in a tree-structure intended to describe an hierarchy of stored keys or indexes, and not processor nodes in a distributed data-base.
10

It is also known in the case of logically sequential data sets to store solely the difference information between two mutually sequential data sets.
15

This procedure is used, for instance, in image processing where a large quantity of information is required to store and reproduce moving pictures or images, but where the difference between two mutually sequential images may be very small, and consequently only the difference between these images is stored and utilised when reproducing the images.
20

It is also known to store solely the difference between two interlinked elements instead of storing the entire element, in conjunction with linked lists.
25

This requires less storage space, but provides the same information.

30 It should also be mentioned that in conjunction with file names or http addresses for instance, different formats and suffixes are used which can make it difficult to describe the difference information between two adjacent nodes in a simple manner.

35

The following publications describe the handling of information structured in a B-tree.

"Transaction Processing: Concepts and Techniques", J. Gray, A. Reuter, Morgan Kaufman, 1993.

5 "Prefix B-trees", R. Bayer, K. Unterauer, ACM Trans. on Database Systems, Vol. 2, No. 1, pp. 11-26, 1977.

10 It should also be mentioned that it is known to store different elements in a so-called container, where the elements lie consecutively one after the other, and where a container can be treated as a unit.

SUMMARY OF THE INVENTION

TECHNICAL PROBLEMS

15

When considering the state of the prior art as described above, it will be seen that a technical problem resides in providing a structure for the storage of data-information in a data-base which requires far less storage space than that
20 required with known structures, even with respect to known compression, by storing solely difference information in relation to preceding nodes.

25 Another technical problem resides in the provision of a compression structure that will also enable a desired element or object to be found effectively.

30 Another technical problem resides in the ability to provide a structure in which a data-word is able to describe the occurring differences in an unequivocal and simple fashion.

Yet another technical problem resides in providing a structure for a data-word which handles different suffixes of the data-information to be stored in a simple fashion.

35

Another technical problem resides in enabling a difference between two sets of data-information to be easily represented, where said difference does not lie at the end of the data-

information, to be readily represented without needing to represent those similarities that are found after this difference in data-information.

- 5 Another technical problem resides in the ability to handle changes of elements or objects in a compressed B-tree.

Yet another technical problem is one of enabling elements or objects in a compressed B-tree to be copied.

10

Another technical problem is one of accelerating the search of a B-tree when seeking a specific element.

- 15 Another technical problem is one of providing a structure in which compression of data-information provides the possibility of substantially speeding-up the search of a B-tree when seeking a specific element.

- 20 Still another technical problem is one of providing a structure which will enable elements to be inserted in and removed from a B-tree simply and effectively, where the data-information is compressed and includes solely difference information relating to preceding nodes.

- 25 Another technical problem is one of providing a structure whereby a node that includes a plurality of elements is available by a read-in to the cache memory.

- 30 Another technical problem is one of providing a structure in which the amount of address information required with references between different nodes and elements included therein is much less than that required traditionally.

SOLUTION

35

With the intention of solving one or more of the aforesaid technical problems and with a starting point from a structure of the kind described in the introduction, it is proposed in

accordance with the invention that the data-information within the root-element constitutes a start-value and that this start-value constitutes the actual data-information. It is also proposed that the differential information within remaining elements is comprised of a data-word that represents the actual difference in the data-information in relation to data-information belonging to preceding elements, regardless of where the difference is within the data-information.

10 Since only the actual difference is stored, regardless of whether this difference is at the end or not, a highly compressed structure is obtained.

15 According to one embodiment of the present invention, such a structure is created by allowing the differential-information to include a directive as to whether something shall be removed relative to the data-information belonging to preceding element and, if so, the nature of what shall be removed and from where in the data-information it shall be removed, and to 20 allow the differential-information to include a directive as to whether something shall be added relative to the data-information belonging to preceding element and, if so, the nature of what shall be added and the place within said data-information in which this addition shall be made.

25 With the intention of providing a compressed or compacted way of expressing this differential-information, it is proposed in accordance with the present invention that a data-word will include a header and an information part, where the header 30 reveals how the information part shall be interpreted.

It is also proposed that the information part can include five different groups of bits, where

35 - a first group which is referred to as a suffix in this document and which represents the type of suffix belonging to the data-information;

- a second group which is referred to as minus information in this document and which represents the number of bytes that shall be removed in relation to preceding element;
- 5 - a third group which is referred to as plus information in this document and which represents the number of bytes that shall be added in relation to preceding element;
- 10 - a fourth group which is referred to as a similarity in this document and which represents the number of bytes from the end, apart from the suffix, that are similar in relation to preceding element; and
- 15 - a fifth group which is designated difference in this document and which constitutes the characters that shall be added, i.e. the plus information, in clear text.

In order to enable the aforesaid groups to be interpreted correctly, it is proposed that the header is comprised of
20 three bits, where

- a first bit shows whether the fourth group exists or not, i.e. whether the difference lies at the end of the data-information or not;
- 25 - a second bit reveals whether or not the second and the third group include four or eight bits; and
- a third bit which reveals whether the second, the third and
30 the fifth groups exist or not, i.e. whether or not any differential-information is found in addition to the suffix.

Thus, there is provided a structure in which the first group includes five bits, the second and third groups include four
35 or eight bits, the fourth group includes zero or eight bits, and the fifth group includes the number of bytes necessary to represent anything that shall be added in relation to preceding element.

Consequently, with the exception of the fifth group, all differential-information can be represented by at least eight bits, i.e. by one byte, when a difference exists only in the suffix, or at the most thirty-two bits, i.e. four bytes when all groups include the maximum number of bits.

One to four bytes plus the actual difference, when this is comprised of plus information, is a significant compression in relation to known techniques. This amount of information requires memory space that corresponds to about 20% of the memory space required with no compression.

Through the medium of its five bits, the first group is able to represent thirty-two different suffixes according to a predetermined list of suffixes, which enables a large number of suffixes to be handled, where each suffix can be represented by only five bits.

The data-information in respective leaf-elements also includes the object sought for.

With the intention of handling locking of objects when an object is the subject of a transaction, it is proposed in accordance with the invention that respective objects include four different bit-combinations, where a first bit-combination represents whether the object is locked or not.

When the first bit-combination shows that an object is unlocked, a second bit-combination constitutes object status bits, a third bit-combination includes information concerning the length of an object key belonging to an object, and a fourth bit-combination includes the object key.

When the first bit-combination shows that the object is locked, the bits belonging to remaining bit-combinations represent a reference to a transaction register. This transaction register includes the object status bits, information con-

cerning the length of an object key belonging to said object, the object key itself, an information concerning the type of lock that has been applied to the object.

- 5 The status bits are utilised in object copying processes, where respective bits indicate whether or not the object has been copied. A plurality of simultaneous copying processes each utilise a respective status bit.
- 10 With the intention of enabling an element to be removed from a B-tree, it is proposed in accordance with the invention that the differential information belonging to a possible following element is adapted to constitute differential information in relation to an element that precedes the removed element, and
- 15 that the address information belonging to a following element is adapted to point to the preceding element, and that the address information belonging to the preceding element is adapted to point to the following element.
- 20 Similarly, when adding an element, it is proposed that the differential information belonging to the added element is adapted to constitute differential information in relation to a preceding element, that the address information belonging to the added element is adapted to point to the preceding element
- 25 and to any following element, that the differential information belonging to said possible following element is adapted to constitute differential information in relation to the added element, that the address information belonging to said following element is adapted to point to the added element,
- 30 and that the address information belonging to the preceding element is adapted to point to the added element.

- With the intention of accelerating search possibilities in a B-tree, it is proposed in accordance with the invention that
- 35 one node is comprised of a container and that respective elements belonging to a node are allocated positions immediately behind one another in said container.

Moreover, the amount of information in a container is adapted to be less than, or equal to, the amount of information that can be read to a cash memory belonging to a processor active within the data-base or utilising the data-base, in one
5 access.

The limitation in the size of a container enables the flexibility afforded by a B-tree to be retained when storing said information in a memory.
10

Despite this size limitation, a container is able to contain a relatively large number of elements, due to the comprehensive compression of respective elements.

15 Thus, when a memory is accessed, a plurality of elements are read into the cache memory at one and the same time. In the case of a normal linked list, it is very often necessary to access the memory for each element that is read-in.

20 Moreover, consecutive storage of the elements in a container requires fewer pointers than required in a linked list, thereby allowing further compression of necessary space.

The possibility of utilising the least possible address information is to place the nodes of the tree-structure on a common
25 page to the greatest possible extent, since references in one page require considerably less address information than references to other pages.

30 ADVANTAGES

One of the primary advantages afforded by an inventive structure is that any differences whatsoever between two adjacent nodes or elements can be represented. When ignoring necessary
35 address information, the storage space required when practicing this method can be reduced to about 20% of the storage space required for a non-compressed B-tree.

Another advantage afforded by the present invention is that a B-tree can be searched with relatively few accesses to the memory or memories in which the information is stored.

5 An inventive tree-structure thus enables optimisation (minimisation) of necessary address storage space in a data structure that is based on linked lists, where the storage structure is an ordered structure that offers a logical, relatively fast search possibility through the linked list.

10

All in all, the invention provides effective handling through the medium of better compression of data than was earlier the case and through a greater possibility of rearranging on one page in conjunction with adding an element to or removing an
15 element from the tree. These rearrangement possibilities are afforded by the fact that each container can be read into the cache memory through one access and handled within the cache memory, which is a very fast and effective handling procedure.

20

The primary characteristic features of an inventive structure are set forth in the characterising clause of the following
Claim 1.

25

BRIEF DESCRIPTION OF THE DRAWINGS

30 So that the invention will be more readily understood and other features thereof made apparent, an inventive structure will now be described with reference to exemplifying embodiments and also with reference to the accompanying drawings, in which

35

Figure 1 is a schematic and highly simplified illustration of a linked list of elements according to known technology;

Figure 2 illustrates part of a linked list according to Figure 1 structured in accordance with a tree-structure in a known manner;

5

Figure 3 illustrates schematically how a data-word intended to represent the differential information between two adjacent elements is structured in accordance with the present invention;

10

Figure 4 is a schematic illustration of the construction of a non-locked object;

Figure 5 is a schematic illustration of a locked object;

15

Figure 6 illustrates schematically how an element can be removed from a tree-structure;

Figure 7 illustrates schematically how an element can be added to a tree-structure;

20

Figure 8 illustrates an alternative embodiment in which the elements belonging to a node are collected in a container;

25

Figure 9 illustrates schematically how a tree can be divided into a number of smaller sub-trees; and

Figure 10 is a schematic illustration of a B-tree and shows how a search through a B-tree can be performed.

30

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

So that the present invention will be more readily understood, a tree-structure intended for use in storing data-information will now be described while using the terminology that will later be used to describe the present invention.

35

A tree-structure represents a method of establishing a linked list in the form of a tree, so as to simplify and make more effective the search for a given element in the list.

- 5 In order to simplify the description of a B-tree, we use an example in which 1000 elements are numbered from 0 to 999. These elements can be placed in a normal linked list A according to Figure 1, or in a tree B according to Figure 2, where a number of elements E1, E2, E3, ... have been distributed in
10 a tree-structure in number order from 0 to 999.

- When searching for elements "312" E313 in a normal link list A, the search is commenced in element "0" E1 which links the search further to element "1" E2, and so on until element
15 "312", E313 is reached. This results in 312 searches, not including the search for the first element "0" E1.

- The tree-structure illustrated in Figure 2 enables the same element to be found by fewer searches.
20

- The illustrated tree has a root-node 1 which includes a root-element 1E that has the value "0-999". This element links to a first intermediate node 2 that includes a plurality of elements 2E1, 2E2, 2E3, ... Each of these elements links to a
25 further intermediate node which includes further elements. Figure 2 shows how element "300-399" 2E4 link to an intermediate node 3.

- The elements 3E1, 3E2, 3E3, ... in the intermediate node 3
30 link, in turn, further to a leaf-nodes that includes a plurality of leaf-elements. Figure 2 shows how the leaf-element "310-319" 3E2 link further to a leaf-nodes 4 that includes a plurality of leaf-elements 4E1, 4E2, 4E3, ...

- 35 The leaf-elements contain the actual information being sought.

In our example, the search is for element "312" 4E3, which in the illustrated tree-structure is carried as follows.

Firstly, a search is made through the elements 2E1, 2E2, 2E3, ... in the first intermediate node 2 until the element "300-399" 2E4 is found, which requires four searches.

5

A search is then made through the second intermediate node 3 until the element "310-319" 3E2 is found, which requires two searches.

10 Finally, a search is made through the leaf-elements 4E1, 4E2, 4E3, ... until the sought element "312" 4E3 is found, which requires three searches.

15 This results in a total of nine searches, not including the search for the root-element "0-999" 1E, as opposed to 312 searches in the case of a normal linked list A.

20 In the case of the normal linked list A, a search for element "999" E1000 requires 1,000 searches, whereas only 30 searches are required in the B-tree B. On the other hand, the elements "0" to "9" are found more quickly in the normal linked list A, since two additional searches are required in the B-tree to come down to the leaf-elements.

25 It is also known in conjunction with linked lists to store solely the difference between two interlinked elements instead of the whole element.

30 Thus, in the Figure 2 example, the root-element 1E might contain data-information "0", all elements in the first intermediate node 2 might then contain data-information "+100", which is the difference in respect of preceding elements, all elements in the second intermediate node 3 might, similarly, contain data-information "+10", and the leaf-elements in the
35 leaf-nodes 4 might contain data-information "+1".

This requires less storage space, but provides the same information.

It is not always that information can be stored as sorted in the simple manner shown in the above example.

- 5 For instance, the data-information is a more complex data structure in the case of file names or http addresses. Moreover, different formats and suffixes are used, which may make it difficult to describe the differential information between two adjacent nodes in a simple manner.

10

The following linked http addresses are given by way of example in this respect:

- 1) <http://www.ericsson.se/ndb/description.html>
15 2) <http://www.ericsson.se/ndb/nodedescription.html>
3) http://www.ericsson.se/ndb/system_arch.pdf
4) http://www.ericsson.se/ndb/system_arch.ps

In this case, the information for 2, 3 and 4 requires 128
20 bytes. This can be described in a compressed form with a starting point from a starting value, which becomes address number 1. In a method in which the difference to preceding elements is stored, it is possible to represent the above addresses in the following way:

25

- 1) <http://www.ericsson.se/ndb/description.html>
2) -"description.html"+"nodedescription.html"
3) -"nodedescription.html"+"system_arch.pdf"
4) -"df"+"s"

30

which requires 80 bytes for 2, 3 and 4 and is thus a compression of the data-information that must be stored in order to reproduce desired information. This corresponds to about 60% of the storage space required in the absence of any compression.
35

The invention essentially builds on a method which is based on solely storing in respective elements the difference to pre-

ceding elements in a B-tree, this method enabling the data-information required in a tree-structure to be further compressed or reduced, in relation to known techniques.

- 5 This method provides the possibility of creating differential information between two adjacent elements without needing to store all of the end of the data-information, provided that the difference does not lie at the end of the element.
- 10 In the above example, the difference between the start-value (1) and the following element (2) is found in a number of characters from the end of the data-information.

This means that the differential information is larger than
15 necessary -"description.html"+"nodedescription.html", since the actual difference between the data-information in the two elements is relatively small in practice, +"node".

It is proposed in accordance with the present invention that
20 the differential information in the various elements is comprised of a data-word that represents the actual difference in the data-information relative to the data-information belonging to preceding element, regardless of where the difference lies in the data-information.

25 This differential information includes a directive as to whether something shall be removed from the data-information belonging to a preceding element and, if such is the case, the nature of what shall be removed and from where in the data-
30 information it shall be removed.

The differential information may also include a directive that something shall be added to the data-information belonging to a preceding element, and, in such case, what shall be added
35 and where in the data-information the addition shall be made.

Figure 3 is intended to show the construction of a data-word. In the illustrated case, a data-word 5 includes a header 51

and an information part 52, where the header 51 discloses how the information part 52 shall be interpreted.

The information part 52 may include five different groups of
5 bits, where

- a first group 521, here designated suffix, represents the type of suffix contained in the data-information;
- 10 - a second group 522, here designated minus information, which represents the number of bytes that shall be removed relative to a preceding element;
- a third group 523, here designated plus information, representing the number of bytes that shall be added to a preceding element;
15
- a fourth group 524, here designated similarity, which represents the number of bytes from the end that are similar to a preceding element, disregarding the suffix; and
20
- a fifth group 525, here designated difference, which constitutes the characters that shall be added, i.e. the plus information 523 in clear text.

25

The header 51 is comprised of three bits, where

- a first bit 511 discloses whether the fourth group 524 exists or not, i.e. whether the difference lies at the end
30 of the data-information or not;
- a second bit 512 which discloses whether the second and third groups 522, 523 include four or eight bits; and
- 35 - a third bit 513 which discloses whether the second, third and fifth groups 522, 523, 525 exist or not, i.e. whether or not any differential information is found with the exception of the suffix 521.

The first group 521 includes five bits, the second and third groups 522, 523 include four or eight bits, the fourth group 524 includes zero or eight bits, and the fifth group 525 includes the number of bytes necessary to represent what shall possibly be added.

Thus, with the exception of the fifth group 525, all differential information can be represented by at least eight bits, that is to say one byte, when a difference lies solely in the suffix, or at most thirty-two bits, that is to say four bytes, when all groups include the maximum number of bits.

Through the medium of its five bits, the first group 521 is able to represent thirty-two different suffixes according to a predetermined suffix list.

For instance, the suffix "html" can be allocated the value "00000" and the suffix "pdf" can be allocated the value "00001". Thus, respective suffixes are represented by 5 bits in both cases. On the other hand, if written-out the suffix "html" requires four bytes, 32 bits, and "pdf" requires three bytes, 24 bits.

To exemplify this, we return to the example of linked http addresses:

- 1) <http://www.ericsson.se/ndb/description.html>
- 2) <http://www.ericsson.se/ndb/nodedescription.html>
- 3) http://www.ericsson.se/ndb/system_arch.pdf
- 4) http://www.ericsson.se/ndb/system_arch.ps

According to the present invention, this can be described in a further compressed form than can be achieved by known techniques, such as:

- 1) <http://www.ericsson.se/ndb/description.html>

- 2) The header 51: 101 (3 bits); suffix 521: html (5 bits);
minus information 522: -0 (4 bits); plus information 523: 4
(4 bits); similarity in the end 524: 11 (8 bits); difference
525: "node" (4 bytes)

5

- 3) The header 51: 001 (3 bits); suffix 521: pdf (5 bits);
minus information 522: -15 (4 bits); plus information 523:
+11 (4 bits); difference 525: "system_arch" (11 bytes)

- 10 4) The header 52: 000 (3 bits) suffix 521: ps (bits)

which results in 21 bytes for rows 2, 3 and 4 as opposed to 80
bytes needed for compression in accordance with known techni-
ques, and 128 bytes in respect of non-compressed information.

15

To this information, there must be added the address informa-
tion required to maintain the tree-structure and to enable a
search to be made through said structure.

- 20 Data-information belonging to respective leaf-elements also
includes the actual object, for instance the internal repre-
sentation for the external http address that a user sees, a
so-called object key.

- 25 An element can be affected by a transaction in different ways,
for instance by a change or updating of the object. In the
case of certain changes, it is necessary to lock the element,
or the object, against other transactions. Different types of
locks can be used, depending on the extent to which the object
30 can be made available for other transactions.

According to the present invention, an object 6 includes four
different bit combinations, shown in Figure 4.

- 35 A first bit combination 61 in the object 6 reveals whether the
object is locked or not.

When the first bit combination 61 reveals that the object 6 is not locked, a second bit combination 62 includes object status bits, a third bit combination 63 includes information concerning the length of an object key belonging to said object, and
5 a fourth bit combination 64 includes the actual object key.

When the first bit combination 61 reveals that the object is locked, the bits belonging to remaining bit combinations represent a reference 65 to a transaction register 7 belonging to
10 the transaction that has caused the lock to be applied, in accordance with Figure 5.

This transaction register 7 includes object status bits 62, information 63 concerning the length of the object key, the
15 object key 64 itself, and information 66 concerning the type of lock that has been applied to the object 6.

The status bits 62 are used in object copying processes, in that respective bits indicate whether the object has been
20 copied or not. Each of a number of simultaneous copying processes utilize its respective status bit, whereby the object is available for copying by as many different copying processes as there are status bits.

Figure 6 is intended to show how an element E'2 is removed in accordance with the present invention. Figure 6 shows that a differential information E'31 belonging to a possible following element E'3 is adapted to constitute difference information in relation to an element E'1 preceding the removed
30 element E'2.

Any address information E'13 belonging to the following element E'3 shall also be adapted to point to the preceding element E'1, and the address information E'13 belonging to the
35 preceding element E'1 shall be adapted to point to the following element E'3.

Figure 7 illustrates in a similar manner that when adding an element E''2, the differential information E''21 belonging to the added element E''2 is adapted to constitute differential information in relation to a preceding element E''1.

5

The differential information E''31 belonging to a possible following element E is three shall then be adapted to constitute differential information in relation to the added element E''2.

10

Any address information E''12, E''23 belonging to the added element E''2 is adapted to point to the preceding element E''1 and the following element E''3, the address information E''23 of any following element E''3 is adapted to point to the added element E''2, and the address information E''12 belonging to the preceding element E''1 is adapted to point to the added element E''2.

15

In the above description, the elements within a node are interlinked in a linked list. This means that a search through these elements may require access to a memory for each element.

20

Since the aforescribed compression structure provides significant compression of necessary memory space, the structure enables a plurality of elements to be stored in a common container.

25

Figure 8 illustrates a preferred embodiment of the present invention where respective nodes are comprised of a container 8.

30

Respective elements 8E1, 8E2, 8E3, ... belonging to a node are allocated positions immediately after one another in the container 8.

35

With the intention of making the search possibilities through an inventive tree-structure still more effective, it is proposed that the amount of information contained in a container 8

is less than, or equal to, the amount of information that can be read through one read-in to a cache memory belonging to a processor that operates in the data-base or utilises said data-base.

5

Because the elements 8E1, 8E2, 8E3, ... in a container 8 lie consecutively one after the other, it is not necessary for respective elements to include all the address information required in a linked list.

10

Thus, when the elements 8E1, 8E2, 8E3, ... are contained in a container, it is proposed in accordance with the invention that the links and associated address information found within respective intermediate elements 8E1, 8E2, 8E3, ... are limited to one link 81 and associated address information concerning a first element 8'E1 belonging to a following node 8' in the hierarchy, and one link 82 and associated address information from respective first intermediate elements 8E1 to an element 8'E3 in a preceding node and associated container 8' in the hierarchy.

20

As similarly, links and associated address information within respective leaf-elements 8'E1, 8'E2, 8'E3, ... are limited to a link 81 and associated address information from respective first leaf-elements 8'E1 to an element 8E3 in a preceding node and associated container 8 in the hierarchy.

25

Different quantities of address information are required when addressing within a table, this quantity depending on how the various nodes are placed. If a node on another page is to be addressed, the addressing procedure may require up to eight bytes in the worst of cases.

30

The pages can be given different sizes and are adapted to different data structure.

35

In one preferred embodiment of the present invention, the tree nodes are placed on one and the same page as far as possible.

This makes it easier to refer to another node. According to the present invention, a container is given a specific size which is adapted to enable its contents to be read into a cache memory through one access.

5

Present-day processors are normally able to read 128 bytes into the cache memory at a time, and the following description is based on this value. It will be understood, however, that the size of a container and other dimensions mentioned in the following description shall be adapted to the amount of information that can be read into the cache memory of the processors concerned at any one time, in order to obtain the desired effect.

15 Another parameter that may need to be taken into account in future applications of the present invention is the time taken to process the information read into the cache memory. When the process time exceeds the time for a cache-miss, it is suitable to limit the size of a container, and other dimensions, to a value at which the time taken to process a container is shorter than the time taken for a cache-miss regardless of the amount of information that can be read into the cache memory at a time.

25 When wishing to address all containers that obtain room on one page with an adapted small amount of address information, it is proposed in accordance with the invention that one page is adapted to at least accommodate the number of containers that can be addressed with the aid of one byte. 256 containers can be addressed with one byte, which corresponds to 32,768 bytes. Thus, one page shall be adapted to be able to include at least 33 Kbytes and other necessary information belonging to one page.

35 Figure 9 shows schematically how a B-tree has been divided into a number of smaller part-trees T1, T2, T3. This enables a tree which is too large to be accommodated on one page to be divided into smaller part-trees.

Where a part-tree T1 is accommodated on one page p1, the necessary address information R1 requires only one byte per reference. When the references R2, R3 are between the part-trees T1, T2, T3, a larger quantity of address information is required, namely the quantity of information required to address containers on different pages p1, p2, ..., pn.

Thus, the majority of references in the total tree can be accommodated within one byte each. Solely those references between the part-trees require a larger memory space.

One method of searching through a B-tree in accordance with the present invention will now be described by way of example.

Figure 10 is a schematic illustration of a tree. A search begins from a so-called root-page and continues downwards. Since the elements of the B-tree contain solely differential information, it is necessary to have a start-value at all search levels. The actual object that is stored can be found by utilising the start-value and the differential information. The root-node on the root-page contains the start-value "NULL".

The desired object is compared with a first element 9e11. If the object is smaller than or equal to the element 9e11, the search follows the left pointer 9p11, otherwise there follows a comparison with the next element 9e12. If the object is smaller than or equal to this element 9e12, the search follows the left pointer 9p12, which is the case in our example.

The start-value for the node to which the pointer 9p12 points is then the element 9e12, since all elements within this node are greater than the element 9e12. The search continues in this way, down through the nodes on one page.

If the search passes to another page, this page will have the same start-value as the root-node as though the search had

continued through the tree without paying attention to the pages. It will be evident from the above that the start-value is known implicitly throughout the entire search. Thus, it is not necessary to store the start-values in conjunction with
5 searches.

If an object cannot be accommodated in a container, a further container is allocated and combined with the first container, thus forming a larger container that has the necessary space.
10

This container is then too large for its contents to be read into a cache memory on one and the same occasion, which results in a cache-miss in conjunction with a search for certain elements within this container. This provides a functionality
15 which, however, is paid for by a loss in performance.

It will be understood that the invention is not restricted to the aforescribed and illustrated exemplifying embodiments thereof and that modifications and variations can be made
20 within the scope of the inventive concept as illustrated in the following Claims.

CLAIMS

1. A structure for the storage of data-information within a data-base, in which data-information is structured in a node hierarchy that comprises a tree-structure that includes a root-node, possibly one or more levels of intermediate nodes, and one or more leaf-nodes, wherein the root-node includes one or more root-elements, respective intermediate nodes include one or more intermediate elements, and respective leaf-nodes include one or more leaf-elements, wherein a first intermediate element within an intermediate node is linked to the root-element or to an intermediate element within an intermediate node placed at a higher level in the hierarchy, and wherein remaining intermediate elements within the intermediate node are linked to said first intermediate element, wherein a first leaf-element within respective leaf-nodes is linked either to the root-element or to an intermediate element of one of the intermediate nodes and remaining leaf-elements in a leaf-nodes are linked to the first leaf-element, wherein respective elements include address-information for the links and the data-information, where the order of the elements within a node is a logical order with respect to data-information, such as an alphabetical or numerical order, and where the data-information in respective elements constitutes the difference in data-information in relation to the data-information belonging to preceding element in the hierarchy, c h a r a c t e r i s e d in that the data-information within said root-element constitutes a start-value; in that said start-value is comprised of the actual data-information; in that differential information within remaining elements is comprised of a data-word that represents the actual difference of the data-information in relation to the data-information that belongs to the preceding element, regardless of where the difference lies within said data-information.

2. A structure according to Claim 1, c h a r a c -
t e r i s e d in that said differential information

includes information as to whether something shall be removed relative to the data-information belonging to the preceding element and, in such case, what shall be removed and from where within said data-information it shall be removed; and in
5 that said differential information includes a directive as to whether something shall be added relative to the data-information belonging to the preceding element and, in such case, what shall be added and where within said data-information the addition shall be made.

10

3. A structure according to Claim 2, c h a r a c -
t e r i s e d in that said data-word includes a header
and an information part; and in that said header reveals how
said information part shall be interpreted.

15

4. A structure according to Claim 3, c h a r a c -
t e r i s e d in that said information part may include
five different groups of bits, wherein

20 - a first group, here designated suffix, represents the type
of suffix possessed by said data-information;

- a second group, here designated minus information, which
represents the number of bytes that shall be removed
25 relative to the preceding element;

- a third group, here designated plus information, which
represents the number of bytes that shall be added in
relation to the preceding element;

30

- a fourth group, here designated similarity, which represents
the number of bytes from the end that have similarity in
relation to the preceding element, apart from said suffix;
and

35

- a fifth group, here designated difference, which constitutes
the characters that shall be added to the data-information,
i.e. said plus information, in clear text.

5. A structure according to Claim 4, c h a r a c -
t e r i s e d in that said header is comprised of three
bits, wherein

5

- a first bit reveals whether or not the fourth group exists,
i.e. whether or not the difference lies at the end of said
data-information;

10 - a second bit reveals whether said second and third groups
include four or eight bits; and wherein

- a third bit reveals whether or not said second, third and
fourth groups exists, i.e. whether or not differential
15 information exists apart from said suffix.

6. A structure according to Claim 5, c h a r a c -
t e r i s e d in that said first group includes five
bits; in that said second and third groups include four or
20 eight bits; in that said fourth group includes zero or eight
bits; and in that said fifth group includes the necessary
number of bytes.

7. A structure according to Claim 6, c h a r a c -
25 t e r i s e d in that said first group is able to repre-
sent thirty-two different suffixes in accordance with a prede-
termined list of suffixes, through the medium of its five
bits.

30 8. A structure according to Claim 8, c h a r a c -
t e r i s e d in that for respective leaf-elements said
data-information includes a sought object; in that said object
includes four different bit-combinations where a first bit
combination represents whether the object is locked or not.

35

9. A structure according to Claim 8, c h a r a c -
t e r i s e d in that when said first bit combination
shows that said object is not locked, a second bit combination

includes object status bits, a third bit combination includes information concerning the length of an object key belonging to said object, and a fourth bit combination includes said object key.

5

10. A structure according to Claim 8, c h a r a c -
t e r i s e d in that when said first bit combination
reveals that said object is locked, the bits belonging to
remaining bit combinations represent a reference to a trans-
action register; in that said transaction register includes
object status bits, information concerning the length of an
object key belonging to said object, said object key itself,
and information concerning the type of lock that has been
applied to said object.

15

11. A structure according to Claims 9 and 10, c h a r -
a c t e r i s e d in that said status bits are used in
object copying processes; in that respective bits indicate
whether the object has been copied or not; and in that a
plurality of simultaneous copying processes each utilise a
respective status bit.

12. A structure according to Claim 1, c h a r a c -
t e r i s e d in that when removing an element, the dif-
ferential information belonging to a possible following ele-
ment is adapted to constitute differential information rela-
tive to an element that precedes the removed element; in that
address information belonging to said following element is
adapted to point to said preceding element; and in that the
address information belonging to said preceding element is
adapted to point to said following element.

13. A structure according to Claim 1, c h a r a c -
t e r i s e d in that when adding an element, the dif-
ferential information belonging to said added element is adap-
ted to constitute differential information relative to a pre-
ceding element; in that the address information belonging to
said added element is adapted to point to said preceding ele-

ment and a possible following element; in that differential information belonging to said possible following element is adapted to constitute differential information relative to said added element; in that the address information belonging to said following element is adapted to point to said added element; and in that address information belonging to said preceding element is adapted to point to said added element.

14. A structure according to Claim 1, c h a r a c -
t e r i s e d in that a node is comprised of a container;
and in that respective elements belonging to a node are allocated immediately sequential positions within said container.

15. A structure according to Claim 14, c h a r a c -
t e r i s e d in that a container contains less information than, or the same amount of information as that amount of information that can be read into a cache memory belonging to a processor acting within said data-base or using said data-base, at one and the same time.

16. A structure according to Claim 14 or 15, c h a r -
a c t e r i s e d in that said links and associated address information within respective intermediate elements are limited to a link and its associated address information concerning a first element belonging to a following node in said hierarchy; and in that said links and their associated address information within respective leaf-elements are limited to one link and its associated address information from respective first leaf-elements to an element belonging to a preceding node in said hierarchy.

17. A structure according to Claim 13, c h a r a c -
t e r i s e d in that one page is adapted to accommodate a plurality of nodes that can be addressed by means of one byte; in that a tree-structure is divided into a plurality of part-trees where respective part-trees are included by one page; in that address information belonging to respective references between two nodes within a common part-tree is

comprised of one byte; and in that the address information belonging to respective references between two different part-trees is comprised of the memory space necessary therefor.

1/3

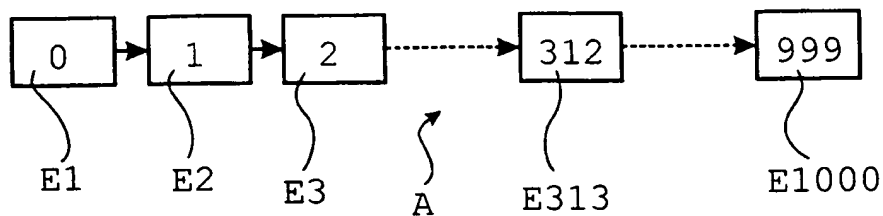


Fig. 1.

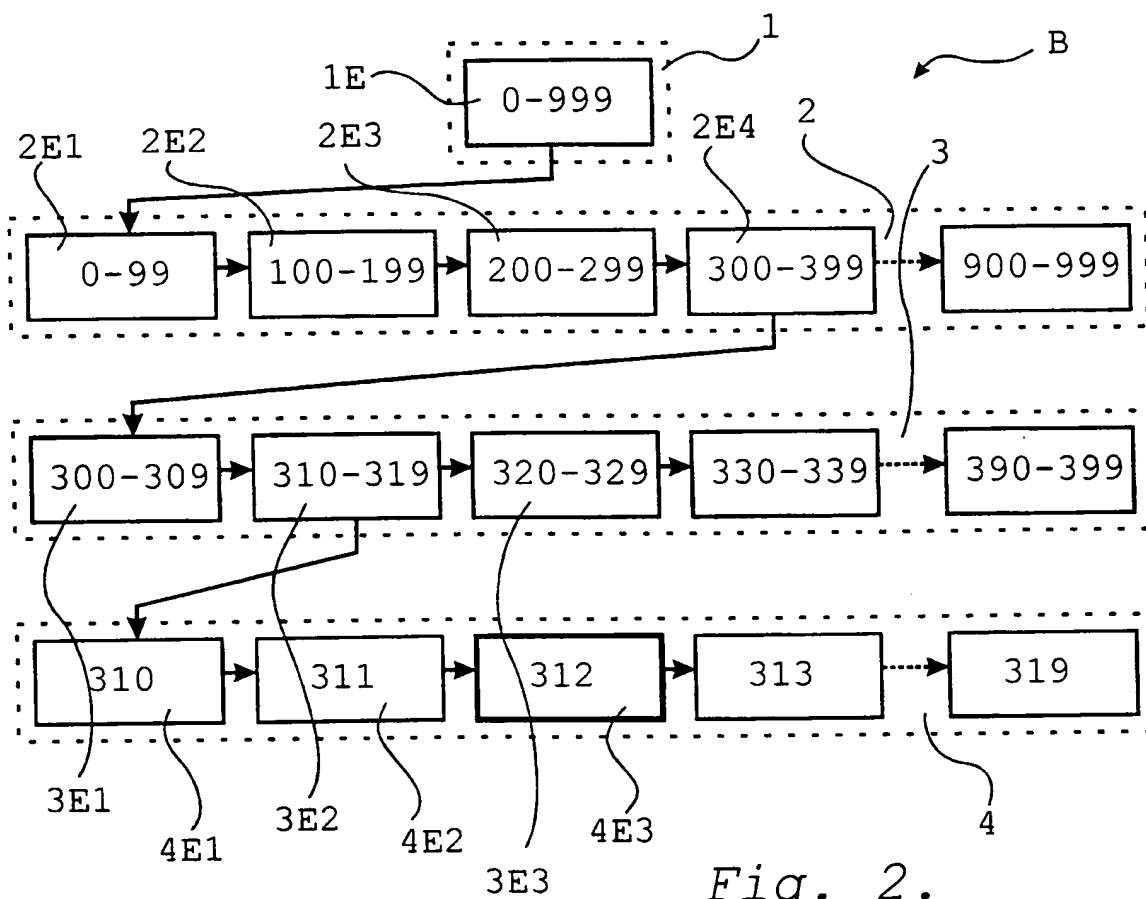


Fig. 2.

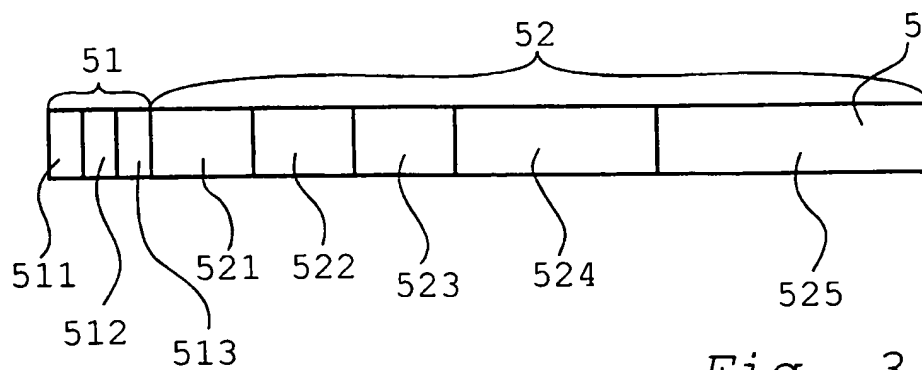


Fig. 3.

2/3

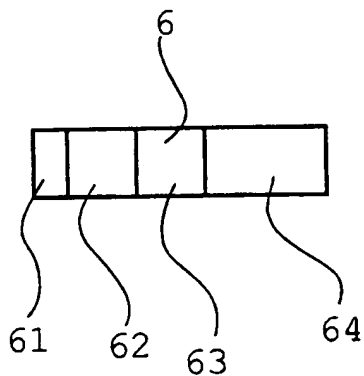


Fig. 4.

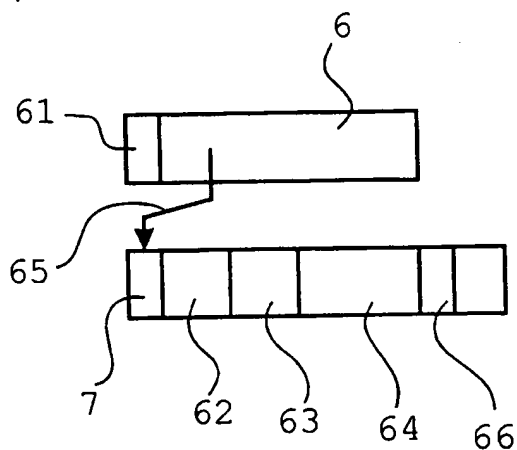


Fig. 5.

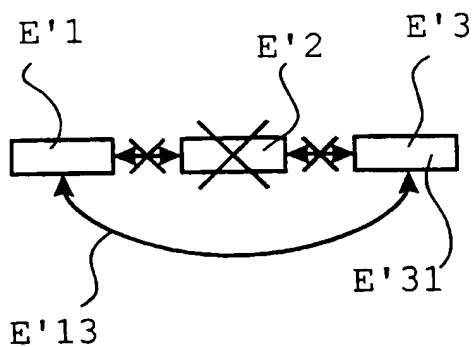


Fig. 6.

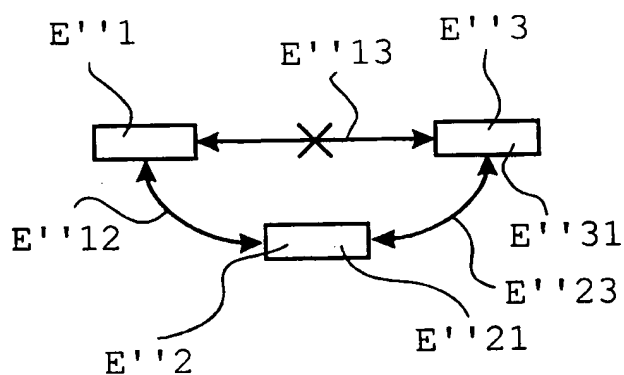


Fig. 7.

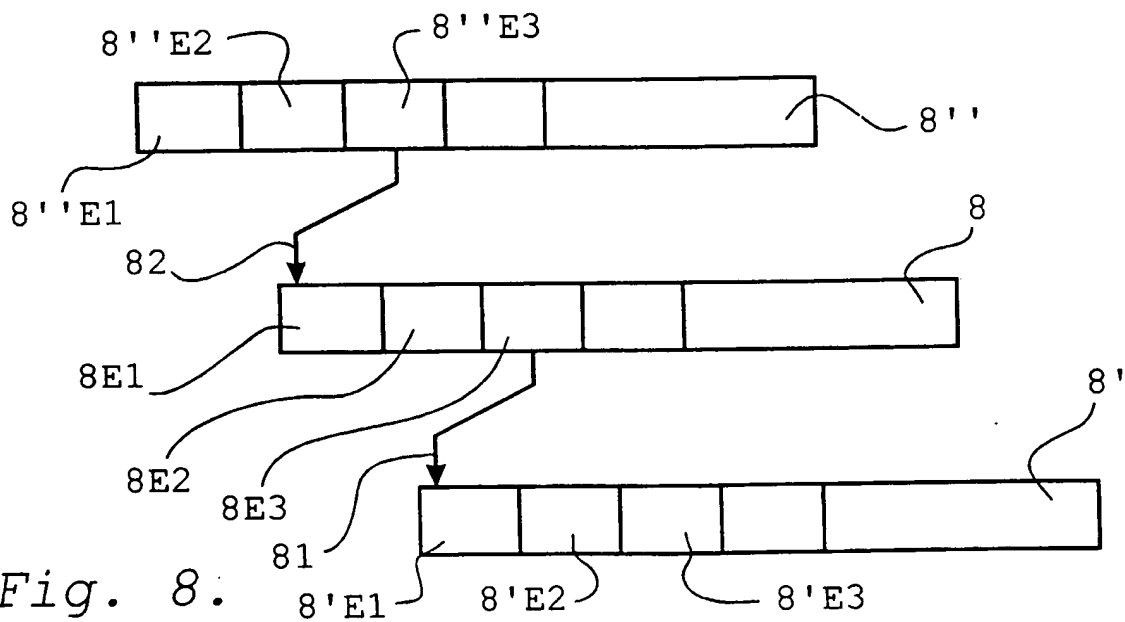


Fig. 8.

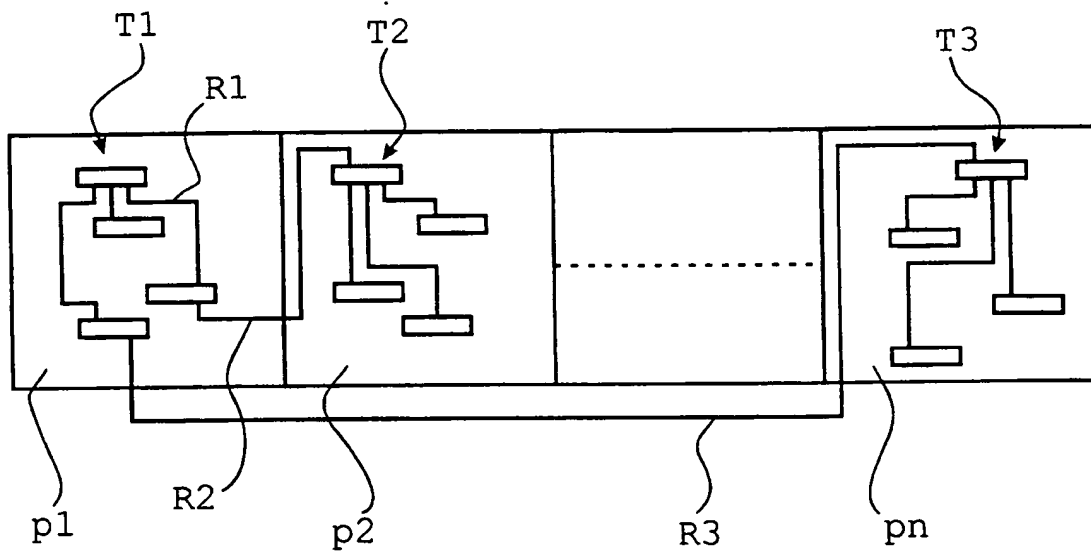


Fig. 9.

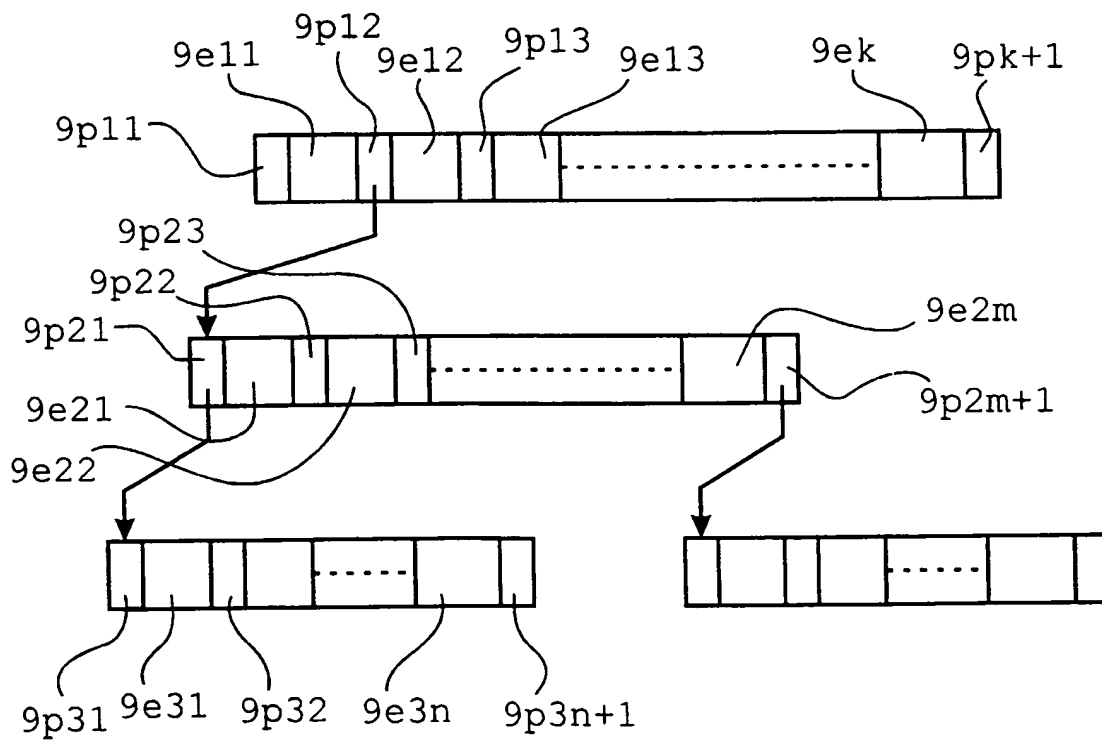


Fig. 10.